**Homework 5 Log**

**Assumptions:**

1. Removed all rows from auto-data.txt with an ‘NA’.
2. Decision tree does not use majority voting. Instead, it uses a probability.
3. If we pruned, we would say that there is a 50/50 chance of males in second class surviving.
4. We would also say that any males in the crew and any females in first class would have a 100% chance of death.

**Results/Issues:**

1. Found/fixed a bug from our HW4.
2. When the probability was zero, then python would throw an error because it’s out of range for the log function.
3. Enew was returning some values greater than 2 for the auto-data set. We still just chose the smallest Enew.
4. Decision trees worked well for the titanic data set because it has so many rows, but didn’t work well for the auto data because that data set does not have many rows.

SURVIVED yes no Total Recognition (%)

---------- ----- ---- ------- -----------------

yes 1115 375 1490 74.8322

no 328 383 711 53.8678

MPG 1 2 3 4 5 6 7 8 9 10 Total Recognition (%)

----- --- --- --- --- --- --- --- --- --- ---- ------- -----------------

1 4 1 2 5 5 4 3 1 1 0 26 15.3846

2 4 2 2 6 4 3 2 3 0 0 26 7.69231

3 3 3 3 5 4 2 3 2 0 0 25 12

4 2 5 3 4 5 1 4 1 0 0 25 16

5 2 4 2 7 5 2 2 1 0 0 25 20

6 1 5 2 5 5 2 3 2 0 0 25 8

7 1 2 4 7 8 0 1 2 0 0 25 4

8 2 5 2 3 4 5 2 2 0 0 25 8

9 2 3 3 8 3 2 2 2 0 0 25 0

10 3 4 2 5 2 6 3 1 0 0 26 0

IF attr2 == male AND attr0 == second AND attr1 == adult THEN yes: 53.9% no: 46.1%

IF attr2 == male AND attr0 == second AND attr1 == child THEN yes: 54.8% no: 45.2%

IF attr2 == male AND attr0 == crew AND attr1 == adult THEN no: 97.2% yes: 2.8%

IF attr2 == male AND attr0 == crew AND attr1 == child THEN no: 100.0%

IF attr2 == male AND attr0 == third AND attr1 == adult THEN no: 87.0% yes: 13.0%

IF attr2 == male AND attr0 == first AND attr1 == adult THEN no: 86.0% yes: 14.0%

IF attr2 == male AND attr0 == first AND attr1 == child THEN no: 100.0%

IF attr2 == female AND attr0 == second AND attr1 == adult THEN yes: 83.8% no: 16.2%

IF attr2 == female AND attr0 == second AND attr1 == child THEN yes: 72.9% no: 27.1%

IF attr2 == female AND attr0 == first AND attr1 == adult THEN yes: 91.7% no: 8.3%

IF attr2 == female AND attr0 == first AND attr1 == child THEN no: 100.0%

IF attr2 == female AND attr0 == third AND attr1 == adult THEN no: 22.3% yes: 77.7%

IF attr2 == female AND attr0 == crew AND attr1 == adult THEN yes: 67.4% no: 32.6%

IF attr2 == female AND attr0 == crew AND attr1 == child THEN no: 100.0%

IF attr4 == 1 AND attr6 == 77 AND attr1 == 4 THEN 8: 33.3% 7: 66.7%

IF attr4 == 1 AND attr6 == 76 AND attr1 == 4 THEN 7: 66.7% 8: 33.3%

IF attr4 == 1 AND attr6 == 75 AND attr1 == 4 THEN 7: 50.0% 8: 50.0%

IF attr4 == 1 AND attr6 == 74 AND attr1 == 4 THEN 8: 33.3% 6: 33.3% 7: 33.3%

IF attr4 == 1 AND attr6 == 73 AND attr1 == 4 THEN 6: 50.0% 7: 50.0%

IF attr4 == 1 AND attr6 == 71 AND attr1 == 4 THEN 7: 50.0% 6: 50.0%

IF attr4 == 1 AND attr6 == 79 THEN 8: 100.0%

IF attr4 == 1 AND attr6 == 78 AND attr1 == 4 THEN 9: 25.0% 8: 75.0%

IF attr4 == 3 AND attr6 == 77 AND attr1 == 3 THEN 5: 100.0%

IF attr4 == 3 AND attr6 == 77 AND attr1 == 4 THEN 6: 66.7% 5: 33.3%

IF attr4 == 3 AND attr6 == 76 THEN 6: 100.0%

IF attr4 == 3 AND attr6 == 75 AND attr1 == 4 THEN 5: 85.7% 6: 14.3%

IF attr4 == 3 AND attr6 == 75 AND attr1 == 6 THEN 4: 100.0%

IF attr4 == 3 AND attr6 == 74 AND attr1 == 4 THEN 6: 100.0%

IF attr4 == 3 AND attr6 == 74 AND attr1 == 6 THEN 4: 100.0%

IF attr4 == 3 AND attr6 == 73 AND attr1 == 4 THEN 4: 66.7% 5: 33.3%

IF attr4 == 3 AND attr6 == 73 AND attr1 == 6 THEN 4: 100.0%

IF attr4 == 3 AND attr6 == 72 AND attr1 == 4 THEN 4: 25.0% 5: 75.0%

IF attr4 == 3 AND attr6 == 71 THEN 4: 100.0%

IF attr4 == 3 AND attr6 == 70 THEN 5: 100.0%

IF attr4 == 3 AND attr6 == 79 AND attr1 == 4 THEN 5: 25.0% 7: 50.0% 8: 25.0%

IF attr4 == 3 AND attr6 == 79 AND attr1 == 6 THEN 4: 50.0% 6: 50.0%

IF attr4 == 3 AND attr6 == 78 AND attr1 == 5 THEN 5: 100.0%

IF attr4 == 3 AND attr6 == 78 AND attr1 == 4 THEN 6: 16.7% 7: 16.7% 5: 66.7%

IF attr4 == 3 AND attr6 == 78 AND attr1 == 6 THEN 5: 100.0%

IF attr4 == 2 AND attr6 == 77 AND attr1 == 4 THEN 8: 33.3% 6: 33.3% 7: 33.3%

IF attr4 == 2 AND attr6 == 76 AND attr1 == 4 THEN 7: 60.0% 6: 40.0%

IF attr4 == 2 AND attr6 == 75 AND attr1 == 4 THEN 7: 50.0% 6: 50.0%

IF attr4 == 2 AND attr6 == 74 AND attr1 == 4 THEN 6: 44.4% 7: 33.3% 5: 22.2%

IF attr4 == 2 AND attr6 == 73 AND attr1 == 3 THEN 4: 100.0%

IF attr4 == 2 AND attr6 == 73 AND attr1 == 4 THEN 4: 50.0% 5: 25.0% 6: 25.0%

IF attr4 == 2 AND attr6 == 72 AND attr1 == 3 THEN 4: 100.0%

IF attr4 == 2 AND attr6 == 72 AND attr1 == 4 THEN 5: 44.4% 6: 33.3% 4: 11.1% 7: 11.1%

IF attr4 == 2 AND attr6 == 71 AND attr1 == 4 THEN 7: 60.0% 6: 20.0% 5: 20.0%

IF attr4 == 2 AND attr6 == 70 THEN 5: 100.0%

IF attr4 == 2 AND attr6 == 79 AND attr1 == 4 THEN 8: 66.7% 9: 33.3%

IF attr4 == 2 AND attr6 == 78 THEN 7: 100.0%

IF attr4 == 5 AND attr6 == 77 AND attr1 == 8 THEN 4: 12.5% 3: 75.0% 2: 12.5%

IF attr4 == 5 AND attr6 == 77 AND attr1 == 6 THEN 4: 100.0%

IF attr4 == 5 AND attr6 == 76 AND attr1 == 8 THEN 4: 14.3% 3: 42.9% 1: 42.9%

IF attr4 == 5 AND attr6 == 76 AND attr1 == 6 THEN 4: 75.0% 3: 25.0%

IF attr4 == 5 AND attr6 == 75 AND attr1 == 8 THEN 3: 50.0% 2: 50.0%

IF attr4 == 5 AND attr6 == 75 AND attr1 == 6 THEN 3: 33.3% 2: 33.3% 4: 33.3%

IF attr4 == 5 AND attr6 == 74 AND attr1 == 8 THEN 3: 20.0% 1: 20.0% 2: 60.0%

IF attr4 == 5 AND attr6 == 74 AND attr1 == 6 THEN 3: 50.0% 4: 50.0%

IF attr4 == 5 AND attr6 == 73 AND attr1 == 8 THEN 1: 64.7% 2: 29.4% 3: 5.9%

IF attr4 == 5 AND attr6 == 72 AND attr1 == 8 THEN 1: 54.5% 2: 36.4% 3: 9.1%

IF attr4 == 5 AND attr6 == 71 AND attr1 == 8 THEN 2: 50.0% 1: 50.0%

IF attr4 == 5 AND attr6 == 70 AND attr1 == 8 THEN 4: 11.1% 2: 88.9%

IF attr4 == 5 AND attr6 == 79 AND attr1 == 8 THEN 3: 44.4% 4: 44.4% 5: 11.1%

IF attr4 == 5 AND attr6 == 79 AND attr1 == 5 THEN 6: 100.0%

IF attr4 == 5 AND attr6 == 78 AND attr1 == 8 THEN 4: 66.7% 5: 33.3%

IF attr4 == 5 AND attr6 == 78 AND attr1 == 6 THEN 4: 100.0%

IF attr4 == 4 AND attr6 == 77 THEN 5: 100.0%

IF attr4 == 4 AND attr6 == 76 AND attr1 == 4 THEN 4: 100.0%

IF attr4 == 4 AND attr6 == 76 AND attr1 == 6 THEN 5: 80.0% 4: 20.0%

IF attr4 == 4 AND attr6 == 75 AND attr1 == 8 THEN 4: 50.0% 1: 50.0%

IF attr4 == 4 AND attr6 == 75 AND attr1 == 6 THEN 4: 50.0% 2: 33.3% 5: 16.7%

IF attr4 == 4 AND attr6 == 74 AND attr1 == 6 THEN 4: 50.0% 2: 50.0%

IF attr4 == 4 AND attr6 == 73 AND attr1 == 8 THEN 2: 100.0%

IF attr4 == 4 AND attr6 == 73 AND attr1 == 6 THEN 4: 66.7% 3: 33.3%

IF attr4 == 4 AND attr6 == 71 AND attr1 == 6 THEN 3: 20.0% 4: 80.0%

IF attr4 == 4 AND attr6 == 70 AND attr1 == 8 THEN 4: 33.3% 3: 33.3% 2: 33.3%

IF attr4 == 4 AND attr6 == 79 AND attr1 == 8 THEN 5: 100.0%

IF attr4 == 4 AND attr6 == 79 AND attr1 == 4 THEN 7: 100.0%

IF attr4 == 4 AND attr6 == 79 AND attr1 == 6 THEN 5: 100.0%

IF attr4 == 4 AND attr6 == 78 AND attr1 == 8 THEN 4: 100.0%

IF attr4 == 4 AND attr6 == 78 AND attr1 == 6 THEN 5: 44.4% 4: 33.3% 3: 22.2%